PCT/US04/20035

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68

A2

(11) International Publication Number: **WO 99/57315**

(43) International Publication Date: 11 November 1999 (11.11.99)

(21) International Application Number: PCT/GB99/01447

(22) International Filing Date: 7 May 1999 (07.05.99)

(30) Priority Data:
9809764.5 7 May 1998 (07.05.98) GB

(71) Applicant *(for all designated States except US)*: ISIS INNOVATION LIMITED [GB/GB]; 2 South Parks Road, Oxford, Oxfordshire OX1 3UB (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ZHANG, Bai, Ping [CN/GB]; 35 Cramner Drive, Southampton, Hampshire SO16 0TD (GB). YE, Shu [GB/GB]; 35 Cramner Drive, Southampton, Hampshire SO16 0TD (GB). HENNEY, Adriano [GB/GB]; 34 Ploughmans Way, Tytherington, Macclesfield, Cheshire SK10 2UN (GB).

(74) Agents: PENNANT, Pyers et al.; Stevens, Hewlett & Perkins, 1 Serjeants' Inn, Fleet Street, London, Greater London EC4Y 1NT (GB).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
*Without international search report and to be republished upon receipt of that report.*



(54) Title: MMP–9 GENE POLYMORPHISMS

(57) Abstract

A method of diagnosis or prognosis of a disease or characterisation of a genotype or genetic sequence which method comprises providing a nucleic acid sample and determining the nature of one or more genetic variants of an MMP–9 gene. Also described are oligonucleotides which hybridise to variants of a polymorphic sequence in the MMP–9 gene or its reverse complement. The method and oligonucleotides are of particular use for the diagnosis and prognosis of diseases characterised by metalloproteinase mediated remodelling, such as atherosclerosis, tumour invasion and metastasis, or inflammatory disease.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

# MMP-9 GENE POLYMORPHISMS

## Introduction

Atherosclerosis, the most common form of arterial disease in Western societies, is caused by the deposition of lipid, migration and proliferation of macrophages and smooth muscle cells, and accumulation of extracellular matrix proteins to form plaques which restrict blood flow. This process, termed vascular remodelling, is dependent for its extent and content on the amount and distribution of extracellular matrix. Several studies in the last few years suggest that matrix metalloproteinases (MMPs) may be responsible, by their expression levels, for the equilibrium between synthesis and degradation of extracellular matrix proteins and thus progression of diseases which display vascular remodelling.

Matrix metalloproteinases (1, 2) are a family of 16 or more zinc dependent endopeptidases which are generally expressed at low levels in normal adult tissues, but are up-regulated during normal and pathological remodelling processes. In concert, MMPs can degrade all components of the extracellular matrix, each enzyme having its substrate preference. These preferences and the domain structures of MMP enzymes have given rise to three groupings, the interstitial collagenases, the gelatinases and the stromelysins. A new group of membrane-bound MMPs has recently been discovered.

One of the stromelysin group of MMPs, known as stromelysin-1 or MMP-3, is a key member of the MMP family (3). It can degrade types II, IV and IX collagen, proteoglycans, laminin, fibronectin, gelatins and elastin, and can also activate other MMPs. MMP-3 expression is primarily regulated at the level of transcription. A common polymorphism occurs in the MMP-3 promoter sequence, with one allele having a run of six adenosines (6A) and the other five adenosines (5A). In transient transfection experiments, a reporter gene under the control of a MMP-3

promoter with 6A at the polymorphic site expresses the reporter gene at a lower level than a construct with the 5A promoter (3). In patients with coronary atherosclerosis, a more rapid progression of the disease is seen in those patients who are homozygous for the 6A allele.

The two gelatinase MMPs so far described, MMP-2 and MMP-9 (also known as Type IV collagenases), share a five-domain structure with the other collagenase and stromelysin MMPs. This five domain structure consists of a signal peptide, a propeptide which contains a cysteine residue that binds to $Zn^{2+}$ at the active site but which is lost upon activation, a catalytic domain that contains the catalytic machinery including the $Zn^{2+}$ binding site, a so-called 'hinge-region' that bridges the catalytic domain and the COOH-terminal domain, and a haemopexin-like COOH-terminal domain thought to be involved in substrate specificity. The gelatinases additionally possess, within their catalytic domain, three fibronectin type II repeats which are thought to facilitate substrate binding.

The promoter regions of MMP genes also share common features. The MMP-9 (gelatinase B) gene (4), found on chromosome 20q11-q13, possesses the polyomavirus enhancer A-binding protein-3 site and the activator protein-1 binding site, which have been found to be important in the regulation of MMP gene expression. The tight control of such expression is important for the balance of tissue turnover. Such control can be exercised at the levels of transcription, activation of latent proenzymes, and inhibition of proteolytic activity by tissue inhibitors of metalloproteinases. Disruption of control at one or more of these levels can give rise to disease caused by over-accumulation or over-degradation of the extracellular matrix. Overproduction of the MMP enzymes leads to the accelerated breakdown of extracellular matrix observed in arthritic disease, tumour invasion and metastasis, whilst inadequate production is associated with excessive accumulation of connective tissue in systemic sclerosis. Recent studies have also indicated a connection between the MMPs and matrix remodelling in the pathogenesis of atherosclerosis (5 - 7).

Thus, MMP activity is associated with inflammatory diseases in general. MMP-9 in particular has been implicated in the pathology of demyelinating neuropathies such as multiple sclerosis (MS) (8). MS is an inflammatory disease of the brain and spinal cord characterised by the formation of focal lesions (plaques) in which demyelination and cell death occurs. The blood brain barrier is damaged in acute lesions, activated lymphocytes and macrophages accumulate and resident glial cells become activated. The infiltrating leukocytes and activated glial cells destroy the myelin sheath which surrounds and insulates the nerve axons. In chronic lesions there is destruction of oligodendrocytes and neurons and the formation of a glial cell scar.

MMP-9 can modulate the passage of leukocytes through the extracellular matrix. It can also degrade a major protein of the myelin sheath called myelin basic protein, producing immunogenic fragments. MMPs such as MMP-9 can process TNF, a potent pro-inflammatory cytokine implicated in the pathogenesis of MS. In both animal models of MS (experimental autoimmune encephalomyelitis – EAEs) and patients with neuroinflammatory disease, activity of MMP-9 is elevated in the cerebrospinal fluid (9, 10). Elevated MMP-9 levels in the blood of patients with MS and other neurological diseases provide a useful indicator of clinical relapse (WO 97/38314).

It has been discovered by the inventors that the nature of certain polymorphic sequences in the promoter, coding sequence and 3' untranslated region of the human MMP-9 gene can affect the severity of atherosclerotic disease. Whereas MMP-3 is associated with progression of disease, MMP-9 is associated with established disease. Determining the nature of one or more of these MMP-9 gene polymorphic sequences, for example determining whether a thymine rather than a cytosine occurs at position -1562 (relative to the start codon) in the promoter of the MMP-9 gene, allows prognosis of the disease. The presence of a thymine rather

than a cytosine nucleotide at position –1562 is associated with a more severe form of disease.

**The Invention**

The invention provides in one aspect, a method of diagnosis or prognosis of a disease or characterisation of a genotype or genetic sequence which method comprises providing a nucleic acid sample and determining the nature of one or more genetic variants of a MMP-9 gene. A MMP-9 gene is herein defined as including the promoter, coding and regulatory sequences.

The invention can be applied to any diseases in which MMP-9 plays a role. These include in particular diseases characterised by metalloproteinase remodelling, such as tumour invasion and metastasis, and inflammatory diseases. The invention is concerned in particular with vascular remodelling diseases such as atherosclerosis, neurological diseases in particular those involving demyelination such as MS, and other conditions in which abnormal breakdown and/or build up of extracellular matrix is a feature, such as arthritic disease and tumour invasion and metastasis.

The genetic variant to be determined may be -1562 cytosine/thymine which is found in the MMP-9 gene promoter at position -1562 bp from the start codon (see Figures 1 and 2). The predominant version of this polymorphism in the population has a cytosine nucleotide present at position -1562, the less common variant version has a thymine nucleotide at position -1562.

The genetic variant to be determined may alternatively be in a sequence encoding a fibronectin-like domain of the MMP-9 enzyme, such as exon 6 +23 guanine/adenine which is found in exon 6 of the MMP-9 coding sequence at position + 23 bp (see Figures 1 and 2). The predominant version of this polymorphism in the population has a guanine

nucleotide present at position +23, the less common variant version has an adenine nucleotide at position +23.

The detection of alternative variants in the MMP-9 promoter coding and regulatory sequence such as -2118 bp (C/T), -1702 bp (T/A), -861bp (C/T), exon 1 +676 bp (C/T), exon 2 +116 bp (G/A), exon 10 +110 bp (C/G), exon 13 +87 bp (G/A), 3' untranslated region +6 bp from stop codon (C/T), and others, is also envisaged in the method of the invention. The above mentioned variants are listed in Table 1.

Table 1

| Positions of Variants | Nature of Variants | Deduced Amino Acid Changes |
|---|---|---|
| *Promoter** | | |
| -2118 bp | C to T | |
| -1702 bp | T to A | |
| -1562 bp | C to T | |
| -861 | C to T | |
| *Coding Region*** | | |
| Exon 1 +676 bp | C to T | Alanine to Valine |
| Exon 2 +116 bp | G to A | Glutamine to Lysine |
| Exon 6 +23 bp | G to A | Arginine to Glutamine |
| Exon 10 +110 bp | C to G | Serine to Cysteine |
| Exon 13 +87 bp | G to A | Valine, conservative |
| *3'-untranslated region* | | |
| 6 bp from beginning of TGA stop codon | C to T | |

* *Positions of promoter variants are calculated from the start of transcription.*

** *Positions of variants in coding region are calculated from the beginning of each respective exon*

In a first embodiment, the method of the invention comprises the steps of: subjecting a region of the MMP-9 gene containing a

polymorphic sequence to a restriction digestion using at least one restriction endonuclease, and analysing the resulting products to determine the presence or absence of a restriction site for the endonuclease as an indication of the nature of the genetic variant.

In the case of the -1562 cytosine/thymine variant, a region of the MMP-9 gene containing the -1562 thymine variant sequence that forms part of a restriction site, can be digested with a restriction enzyme, such as *Bbu*I or *Sph*I, giving rise to a digested product which is then analysed, for example by size determination. A region containing a -1562 cytosine sequence does not possess a restriction site for either *Bbu*I or *Sph*I and will be seen to have not been digested by one of those enzymes upon analysis.

Preferably, the region containing the polymorphic sequence is amplified prior to restriction digestion, for example using a PCR technique.

In an adaption of the first embodiment of the method of the invention, the region of the MMP-9 gene containing a polymorphic sequence is amplified using one or more oligonucleotide primers designed to incorporate into the amplified sequence a restriction endonuclease recognition site comprising a variant of the polymorphic sequence. This is useful where one or other of the variant forms of the polymorphic sequence does not itself contain a restriction site.

Analysis of the restriction digestion products in the method of the invention may be carried out by a size determination technique, for example, using a separation technique such as gel electrophoresis.

In an alternative embodiment of the method of the invention the nature of the variant is determined by using at least one oligonucleotide capable of hybridising specifically to a variant of interest under suitably stringent conditions.

Such an oligonucleotide should be long enough to be specific for a particular variant of interest. Suitably stringent hybridisation conditions need to be employed to enable an oligonucleotide specific for one variant of the polymorphic sequence to hybridise to it and not to any

other variant. This is particularly important when the sequence differences between variants are small, for example –1562 cytosine/thymine. Suitable hybridisation conditions are known and/or can be easily determined. Hybridisation may be detected by means of a detectable label on the specific oligonucleotides.

The specific oligonucleotide may be employed as a sequence specific probe. Methods wherein specific oligonucleotides are employed as primers in an amplification reaction are also provided by the invention. In one such method, the nature of the variant is determined by an amplification reaction which uses one or more sequence specific oligonucleotide primers. A region of the MMP-9 gene containing the polymorphic sequence may be subjected to PCR in the presence of at least one primer designed to hybridise with a particular variant of the polymorphism, at the 3' terminal nucleotide of the primer, such that only fragments comprising that variant will be amplified. Two or more different primers having differing 3' termini may be employed. Oligonucleotides which are complementary to a given DNA sequence except for a mis-match at their 3' termini fail to function as primers in PCR. This technique is known as the amplification refractory mutation system (ARMS) (11).

Other methods of variant sequence detection employing amplification reactions are envisaged. For example, WO 90/11372 describes a method for determining the existence or not of a particular nucleotide at a particular position on a strand of DNA. Primers specific for that nucleotide and having unique tag sequences are employed in an extension reaction. Any resulting extended primers are then immobilised on a support via the tag sequence, for detection. A series of different primers can be used to investigate a number of different alleles at the same time, as long as each different primer is provided with its own unique tag sequence. WO 90/11369 describes a detection method for diagnosis of genetic medical conditions, which involves amplifying a part of the DNA strand of interest using a first pair of liquid phase primers specific to the

target DNA, and then further amplifying using a nested pair of primers, one of which is immobilised or capable of being immobilised.

The principle of the mutagenically separated polymerase chain reaction (MS-PCR) may also be applied in the method according to the invention (12). In an MS-PCR adaptation, the two immobilised primers differ as in ARMS, but are also different to one another in respect of one or more further nucleotides within the terminal few, eg. four, 3' nucleotides. Thus, each primer is mis-matched with its target sequence with respect to one or more nucleotides and those mis-matches differ from mis-matches of the other primer and its target sequence. Advantageously, there may be present in the liquid phase non-immobilised primers corresponding to one or both of the immobilised primers. Such liquid phase primers can act to enhance signal strength and improve discrimination of the primers between their specific target and the other target.

Multiple samples may be subjected to the methods of the invention with the products being analysed by size-determination. A suitable high-throughput method of analysing such products is Microplate Array Diagonal Gel Electrophoresis (MADGE) (13). This system improves the conventional method of genotyping (detection of variant sequences) via PCR followed by allele specific oligonucleotide binding assay in which electrophoresis is used to resolve "bound" from "free" oligonucleotides.

The methods of the invention are not limited to detection of variant sequences which employ size determination by gel-electrophoresis. Various non gel-based tests are known in the art for detecting genetic variants. Much current research in this area utilises "strip technology", in which solid supports with spatially separated probes recognise and trap specific DNA products without the need for gel electrophoresis. The solid support might entrap DNA products resulting from an earlier, separate step such as an amplification reaction, such as herein described. Detection of such products might be enabled through the use of specific oligonucleotides immobilised as parallel lines on membrane based strips.

Amplified targets labeled for example with biotin are hybridised with the immobilised oligonucleotides, and binding may be visualised for example using streptavidin conjugated with alkaline phosphatase and a suitable chromogen.

Also useful in the method according to the invention are detection methods employing arrays of immobilised oligonucleotides, one or more of which has a sequence which is complementary to a variant of the polymorphic sequence. Such arrays are known in the art for example as described in EP373203B. Array techniques are rapid to perform and can detect a number of known polymorphisms simultaneously.

The invention also envisages a combination of tests, which combination includes determining the nature of one or more MMP-9 polymorphisms, for providing an accurate prognosis for remodelling diseases. Such a combination could include for example a test for determining the nature of one or more polymorphisms in other MMP genes such as the MMP-3 gene.

In an alternative aspect, the invention provides an oligonucleotide of between 8 and 50 base pairs, or more preferably 15 to 30 base pairs in length which, under suitably stringent conditions, specifically hybridises to a variant of a polymorphic sequence shown in Table 1, or its reverse complement. This aspect of the invention is concerned in particular with such oligonucleotides which hybridise to the polymorphic sequence as shown in Figure 1, or its reverse complement, in a region containing position -1562. The oligonucleotide may specifically hybridise to the polymorphic sequence containing the variant thymine at position -1562. Particularly preferred is an oligonucleotide which is one of the following sequences, or a fragment of one of the following sequences, including the underlined sequence; and oligonucleotides complementary thereto:

TGGTGGCGCACCTATAATACCA;

TGGTGGCGCATCTATAATACCA;

TGGTGGCGCACCTATAATACCA;

TGGTGGCGCATCTATAATACCA.

In an alternative embodiment, the invention provides an oligonucleotide of between 8 and 50 base pairs, or more preferably 15 to 30 base pairs in length, which under suitably stringent conditions, specifically hybridises to a variant of the polymorphic sequence of Figure 1, or its reverse complement, in a region containing the exon 6 +23 guanine/adenine polymorphism. The oligonucleotide may specifically hybridise to the polymorphic sequence containing the variant adenine at exon 6 +23. Particularly preferred is an oligonucleotide which is one of the following sequences, or a fragment of one of the following sequences including the underlined sequence; and oligonucleotides complementary thereto:

CTCTACACCCGGGACGGCAATG;

CTCTACACCCAGGACGGCAATG;

CTCTACACCCGGGACGGCAATG;

CTCTACACCCAGGACGGCAATG.

The invention also provides oligonucleotide primers having a sequence suitable for the preferential amplification of an MMP-9 DNA template molecule comprising a cytosine nucleotide at position -1562 relative to the MMP-9 gene start codon and oligonucleotide primers having a sequence suitable for the preferential amplification of an MMP-9 DNA template molecule comprising a thymine nucleotide at position -1562.

Further provided are oligonucleotide primers suitable for preferentially amplifying a sequence with a guanine nucleotide at position

+23 within exon 6 of the MMP-9 gene or primers preferentially amplifying a sequence with adenine at position +23 within exon 6 of the MMP-9 gene.

The invention also relates to proteins or peptides encoded by the herein described variant sequences of the MMP-9 gene, produced by methods well known in the art. Specific antibodies raised against such proteins or peptides are also envisaged. Proteins or peptides encoded by the variant sequences may be utilised for screening compounds that bind specifically to a molecule encoded by one variant of a polymorphic sequence, in search for compounds which modulate the activity of the enzyme. The structure of such variant proteins may be used for rational drug design.

18-21

The invention also provides assay kits for use in the detection of MMP-9 gene variants such as those described herein. Such kits may comprise for example one or more oligonucleotides as described herein, and optionally a polymerase enzyme.

The inventors have discovered polymorphisms in the MMP-9 gene sequence including the promoter. The absence or presence of one of certain variant forms of these polymorphisms (-1562 cytosine/thymine) has been found to have a correlation with the severity of atherosclerosis. This information when used alone or in combination with other genetic or clinical data, allows an accurate prognosis of atherosclerosis. The invention will also enable predictions to be made about the expression and/or activity levels of MMP-9 genes and enzymes and hence allow more accurate prognosis of inflammatory diseases and cancer.

The invention is further described in the examples that follow, with reference to the following Figures

Figure 1 shows the genomic promoter and gene sequence of a human MMP-9 gene, including the positions of important polymorphisms.

Figure 2 shows a diagrammatic representation of the Human MMP-9 gene, with exons marked as black boxes and the approximate

position and nature of the polymorphisms discovered by the inventors indicated.

Figure 3 shows the relative activities of a reporter gene under the control of promoters containing MMP-9 variant sequences in transient transfection experiments employing transfected murine lung macrophages (MALU).

## EXAMPLES

**Example 1. Protocol for genotyping:**

A 435bp DNA sequence containing the -1562 (C/T) polymorphic site of MMP-9 was amplified by PCR. The products were then digested with restriction endonuclease *Bbu*I (or *Sph*I) and subjected to agarose gel electrophoresis.

PCR was carried out in a total volume of 25µl containing:

0.1µg of genomic DNA

0.1µg upstream primer (5'-GCCTGGCACATAGTAGGCCC-3')

0.1µg downstream primer (5'-CTTCCTAGCCAGCCGGCATC-3')

200µM each dNTP (dATP, dCTP, dGTP and dTTP)

1× PCR buffer (10mM Tris-HCl, [pH 8.3], 0.001% [w/v] Gelatin, 0.005% W1 [Gibco BRL], 50mM KCL; 1mM $MgCl_2$)

1 unit of Taq DNA polymerase

The solution was overlaid with 50µl mineral oil and incubated for 5 minutes at 95°C, followed by 36 cycles of one minute at 95°C, one minute at 66°C, and two minute at 72°C, with another six minutes at 72°C after the final cycle.

A 20µl aliquot of the PCR reaction was incubated with 2 units of restriction endonuclease *Bbu*I (or *Sph*I) in 1× digestion buffer supplemented with 100µg/ml bovine serum albumin for 16 hours at 37°C.

The digests were then fractionated on a 1.5% agarose gel. Samples of known genotype (verified by sequencing) were used as references, digested with *Bbu*I (*Sph*I) and run alongside the samples being analyzed.

Expected results:

Homozygotes for the C allele (C/C, uncut/uncut): one band - 435bp;

heterozygotes (C/T, uncut/cut): three bands - 435bp, 244bp, and 191bp, respectively;

homozygotes for the T allele (T/T, cut/cut): two bands - 244bp and 191bp respectively.

**Example 2. Association data:**

Five hundred and eighty-four male patients [mean (SD) age: 54.0 (8.1) years] with myocardial infarction and 646 age matched male healthy control subjects [mean (SD): 53.2 (8.4) years], recruited in the ECTIM study (Etude Cas-Témoin de l'Infarctus du Myocarde), a multi-centre, case-control study of myocardial infarction in France and Northern Ireland, were genotyped for the MMP-9 ($C_{-1562}$ -T) polymorphism. There was no significant difference in the frequencies of genotypes or alleles between patients and controls, with the frequency of the T allele being 0.14 in patients and 0.13 in controls.

However, in 421 patients in whom coronary angiographic data were available, there was a significant association between the MMP-9 ($C_{-1562}$-T) polymorphism and severity of coronary atherosclerosis measured by angiography (Table 2): patients carrying one or two copies of the T allele had higher numbers of >50% stenotic arteries than those who were homozygous for the C allele (p=0.033). Thus in 28 percent of patients in the C/T and T/T genotype classes, all three coronary arteries were found to have an atherosclerotic lesion that caused a >50% stenosis. In contrast,

only 16 percent of patients in the C/C genotype class had a >50% stenosis in all three coronary arteries.

Table 2. Number of stenotic (>50%) coronary arteries by genotype

| Genotype | 0 | 1 | 2 | 3 | Total |
|---|---|---|---|---|---|
| C/C | 10 (83.3%) | 151 (75.5%) | 98 (77.8%) | 52 (62.2%) | 311 (73.9%) |
| C/T | 2 (16.7%) | 45 (22.5%) | 27 (21.4%) | 27 (32.5%) | 101 (24.0%) |
| T/T | 0 (-) | 4 (2%) | 1 (0.8%) | 4 (4.8%) | 9 (2.1%) |

*Chi-squared (Mantel-Haenszel) test, $P < 0.05$*

**Example 3. Functional data:**

DNA-protein interaction assays showed that the –1562 bp polymorphic site was located within a DNA sequence recognised by a putative transcription factor, and that the binding affinity was lower when a T (as opposed to a C) was present at the polymorphic site. Along with these observations, an allele-specific effect on transcription was detected by transient transfection experiments in which cultured macrophages were transfected with plasmid constructs containing three concatenated copies of the DNA element (with either a C or a T at the polymorphic site) linked to a minimal promoter upstream of a reporter gene (chloramphenicol acetyl transferase). The DNA elements used for the respective polymorphisms were 'C': 5'-CAGGCGTGGTGGCGCACGCCTATAAT-3' and 'T': 5'-CAGGCGTGGTGGCGCATGCCTATAAT-3'. As shown in Figure 3, a construct with the three copies of DNA element corresponding to the T allele produced two fold more reporter protein than a construct with the

DNA element corresponding to the C allele. These results indicate that the MMP-9 $C_{-1562}$-T is a regulatory functional polymorphism.

**References**

1. Ye S, Humphries S, Henney A. (1998) *Clinical Science* **94**:103-110
2. Birkedal-Hansen H, Moore WGI, Bodden MK, Windsor LJ, Birkedal-Hansen B, DeCarlo A, Engler JA. (1993) *Critical Reviews in Oral Biology and Medicine* **4**:197-250
3. Ye S *et al* (1996) *J. Neuroscience* **271**: 13055-13060
4. Goldberg GI, Marmer BL, Grant GA, Eisen ZA, Wilhelm S, He C. (1989) *Proc. Nat. Acad. Sci USA* **86**:8207-8211
5. Henney AM *et al.* (1991) *Proc. Nat. Acad. Sci USA* **88**:8154-8158
6. Zaltsman AB *et al.* (1997) *Atherosclerosis* **130**:61-70
7. Knox JB *et al.* (1997) *Circulation* **95**:205-212
8. Opdenakker *et al* (1994) *Immunol Today* **15**: 103-107
9. Gijbels K et al (1994) *J.Clin.Invest.*. **94**: 2177-2182
10. Selmaj K *et al* (1991) *Ann. Neurol.* **30**:694-700.
11. Newton CR. *et al* (1989) *Nucl. Acids Res.* **17**:2503-2516
12. Rust S. *et al.* (1993) *Nucl. Acids Res.* **21**:3623-3629
13. Day INM *et al.* (1995) *BioTechniques* **19**:830-835

## CLAIMS

1. A method of diagnosis or prognosis of a disease or characterisation of a genotype or genetic sequence which method comprises providing a nucleic acid sample and determining the nature of one or more genetic variants of a MMP-9 gene.

2. The method as contained in claim 1, wherein the disease is characterised by metalloproteinase mediated remodelling, such as tumour invasion and metastasis, or inflammatory disease.

3. The method as claimed in claim 2, wherein the disease is a vascular disease.

4. The method as claimed in any one of claims 1 to 3 wherein the variant is -1562 cytosine/thymine.

5. The method as claimed in any one of claims 1 to 3 wherein the variant is in a sequence encoding a fibronectin-like domain of the MMP-9 enzyme.

6. The method as claimed in claim 5 wherein the variant is exon 6 +23 guanine/adenine.

7. The method as claimed in any one of claims 1 to 3 wherein the variant or variants are one or more of the variants listed in Table 1.

8. The method as claimed in any one of claims 1 to 7, which method comprises the steps of:

(i) subjecting a region of the MMP-9 gene containing a polymorphic sequence to a restriction digestion using at least one restriction endonuclease; and

(ii) analysing the resulting products to determine the presence or absence of a restriction site for the endonuclease as an indication of the nature of the genetic variant.

9. The method as claimed in claim 8 wherein the variant is -1562 thymine which is in the restriction site that is recognised by the endonuclease whilst -1562 cytosine is not in the restriction site.

10. The method as claimed in claim 9 wherein the restriction endonuclease is *Bbu*I or *Sph*I.

11. The method as claimed in any one of claims 8 to 10 wherein the region of the MMP-9 gene is amplified prior to step (i).

12. The method as claimed in claim 8, which method comprises amplifying a region of the MMP-9 gene containing a polymorphic sequence using one or more oligonucleotide primers designed to incorporate into the amplified sequence a restriction endonuclease recognition site comprising a variant of the polymorphic sequence.

13. The method as claimed in any one of claims 8 to 12 wherein analysis in step (ii) is carried out by a size determination technique.

14. The method as claimed in claim 13 wherein the products are size separated on a gel.

15. The method as claimed in any one of claims 1 to 7 wherein the nature of the variant is determined by using at least one oligonucleotide capable of hybridising specifically to a variant under suitably stringent conditions.

16. The method as claimed in claim 15 wherein the variant of a polymorphic sequence or a sequence complementary thereto, is present in an oligonucleotide immobilised on an array.

17. The method as claimed in claim 15 wherein the oligonucleotide is employed as a primer in an amplification reaction.

18. The method as claimed in any one of claims 1-17 wherein the inflammatory disease is atherosclerosis.

19. An oligonucleotide of between 8 to 50 base pairs in length which, under suitably stringent conditions, specifically hybridises to a variant of a polymorphic sequence shown in Table 1, or its reverse complement.

20. The oligonucleotide as claimed in claim 19 which hybridises to a variant of the polymorphic sequence as shown in Figure 1, or its reverse complement, in a region containing position -1562 .

21. The oligonucleotide as claimed in claim 20 which hybridises to the polymorphic sequence containing the variant thymine at position -1562.

22. The oligonucleotide as claimed in claim 19, which hybridises to a variant of the polymorphic sequence as shown in Figure 1, or its reverse complement, in a region containing the exon 6 +23 guanine/adenine polymorphism.

23. The oligonucleotide as claimed in claim 22, which hybridises to the polymorphic sequence containing the adenine at exon 6 +23.

24. The oligonucleotide as claimed in any one of claims 19 to 23 which is 15 to 30 base pairs in length.

25. An oligonucleotide which is of the following sequences, or a fragment of one of the following sequences, including the underlined sequence; and oligonucleotides complementary thereto:

TGGTGGCGCACCTATAATACCA
TGGTGGCGCATCTATAATACCA
TGGTGGCGCACCTATAATACCA
TGGTGGCGCATCTATAATACCA

26. An oligonucleotide which is of the following sequences, or a fragment of one of the following sequences including the underlined sequence; and oligonucleotides complementary thereto:

CTCTACACCCGGGACGGCAATG
CTCTACACCCAGGACGGCAATG
CTCTACACCCGGGACGGCAATG
CTCTACACCCAGGACGGCAATG

27. An oligonucleotide primer having a sequence suitable for the preferential amplification of an MMP-9 gene DNA template molecule

comprising a cytosine nucleotide at position -1562 relative to the MMP-9 gene start codon.

28. An oligonucleotide primer having a sequence suitable for the preferential amplification of an MMP-9 gene DNA template molecule comprising a thymine nucleotide at position -1562 relative to the MMP-9 gene start codon.

29. An oligonucleotide primer having a sequence suitable for the preferential amplification of an MMP-9 gene DNA template molecule comprising a guanine nucleotide at position +23 within exon 6 of the MMP-9 gene.

30. An oligonucleotide primer having a sequence suitable for the preferential amplification of an MMP-9 gene DNA template molecule comprising a adenine nucleotide at position +23 within exon 6 of the MMP-9 gene.

31. A kit containing at least one oligonucleotide according to any one of claims 20 to 30 and optionally a polymerase enzyme.

32. Use of peptides encoded by the oligonucleotides of claim 19 in a method of screening for compounds which modulate the activity of one or more products of genetic variants of a MMP-9 gene.

## Figure 1

-2174 AAGCTTCAGAGCCAGGCAGTTCTGGGCTTGAACACTAGTTCTGTGGATTAACTCGCTCTGTGATCACAGGCA -2101

-2100 AATTCCTTAACTCTCTGAGCCTTAGTTTCCCCCTCTGAAAACAGGAGGGATACTCATTAAACTTACCTTACAGGTGGTGAGGATGAAACGAGAGGCTTAT -2001

-2000 AGAGAACTTATTACGGTGCTTGACACAGTAAATCTCAAAAAATGCATTATTATTATTATGGTTCAGAGGTAAAGTGACTTGCCCAAGGTCACATAGCTGG -1901

-1900 AAAATGCAGAGCCGGGATGGAAATCCAGGACTTCGTGACGCAAAGCAGATGTTCATTGGTTAGTGAACTTTAGAACTTCAACTTTTCTGTAAAGGAAGTT -1801

-1800 AATTATCTCCATCTCACAGTCTCATTTATTAGATAAGCATATAAAATGCCTGGCACATAGTAGGCCCTTTAAATACAGCTTATTGGGCCGGGCGCCATGC -1701

-1700 TCATGCCCGTAATCCTAGCACTTTGGGAGGCCAGGTGGGCAGATCACTTGAGTCAGAAGTTCGAAACCAGCCTGGTCAACGTAGTGAAACCCCATCTCTA -1601

T (-1562 polymorphism)

-1600 CTAAAAATACAAAAAATTTAGCCAGGCGTGGTGGCGCACCTATAATACCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGAT -1501

-1500 GTTGCAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAGTGATACTACCCCCCCCAAAAATAAAATAAAATAAATAAATACAACTTTTT -1401

-1400 GAGTTGTTAGCAGGTTTTTCCCAAATAGGGCTTTGAAGAAGGTGAATATAGACCCTGCCCGATGCCGGCTGGCTAGGAAGAAAGGAGTGAGGGAGGCTGC -1301

-1300 TGGTGTGGGAGGCTTGGGAGGGAGGCTTGGCATAAGTGTGATAATTGGGCCTGGAGATTTGGCTGCATGGAGGCAGGGCTGGAGGAACTAAGGGCTCCTA -1201

-1200 TAGATTATTTCCCCATATCCTGCCGCAATTTGCAGTTGAAGAATCCTAAGCTGAGAAAGGGGAGGCATTTACTCCAGGTTACACTGCAGCTTAGAGCCCA -1101

-1100 ATAACCTGGTTTGGTGATTCCAAGTTAGAATCATGGTCTTTTGGCAGGGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGACATAATCATGGCTCACTG -1001

-1000 TATCCTTGACCTTCTTTCTGGGCTCAAGCAATCCTCCCACCTCGGCCTCCCAAAGTGCTAAGATTACAGGAATGAGCCACCATACCTGGCCCTGAATCTT -901

-900 GGGTCTTGGCCTTAGTAATTAAAACCAATCACCACCATCCGTTGCGGACTTACAACCTACAGTGTTCTAAACATTTTATATGTTTGATCTCATTTAATCC -801

-800 TCACATCAATTTAGGGACAAAGAGCCCCCCACCCCCCGTTTTTTTTTTTTACAGCTGAGGAAACACTTCAAAGTGGTAAGACATTTGCCCGAGGTCCTGAA -701

**Figure 1 (continued)**

```
-700 GGAAGAGAGTAAAGCCATGTCTGCTGTTTTCTAGAGGCTGCTACTGTCCCCTTTACTGCCCTGAAGATTCAGCCTGCGGAAGACAGGGGGTTGCCCCAGT -601
-600 GGAATTCCCCAGCCTTGCCTAGCAGAGCCCATTCCTTCCGCCCCCAGATGAAGCAGGGAGAGGAAGCTGAGTCAAAGAAGGCTGTCAGGGAGGGAAAAAG -501
-500 AGGACAGAGCCTGGAGTGTGGGGAGGGGTTTGGGGAGGATATCTGACCTGGGAGGGGGTGTTGCAAAAGGCCAAGGATGGGCCAGGGGGATCATTAGTTT -401
-400 CAGAAAGAAGTCTCAGGGAGTCTTCCATCACTTTCCCTTGGCTGACCACTGGAGGCTTTCAGACCAAGGGATGGGGGATCCCTCCAGCTTCATCCCCCTC -301
-300 CCTCCCTTTCATACAGTTCCCACAAGCTCTGCAGTTTGCAAAACCCTACCCCTCCCCTGAGGGCCTGCGGTTTCCTGCGGGTCTGGGGTCTTGCCTGACT -201
-200 TGGCAGTGGAGACTGCGGGCAGTGGAGAGAGGAGGAGGTGGTGTAAGCCCTTTCTCATGCTGGTGCTGCCACACACACACACACACACACACACACAC -101
-100 ACACACACACACCCTGACCCCTGAGTCAGCACTTGCCTGTCAAGGAGGGGTGGGGTCACAGGAGCGCCTCCTTAAAGCCCCCACAACAGCAGCTGCAGTC -1
+1
AGACACCTCTGCCCTCACCATGAGCCTCTGGCAGCCCCTGGTCCTGGTGCTCCTGGTGCTGGGCTGCTGCTTTGCTGCCCCCAGACAGCGCCAGTCCACC
                    M  S  L  W  Q  P  L  V  L  V  L  L  V  L  G  C  C  F  A  A  P  R  Q  R  Q  S  T
CTTGTGCTCTTCCCTGGAGACCTGAGAACCAATCTCACCGACAGGCAGCTGGCAGAGGAATACCTGTACCGCTATGGTTACACTCGGGTGGCAGAGATGC
L  V  L  F  P  G  D  L  R  T  N  L  T  D  R  Q  L  A  E  E  Y  L  Y  R  Y  G  Y  T  R  V  A  E  M  R
GTGGAGAGTCGAAATCTCTGGGGCCTGCGCTGCTGCTTCTCCAGAAGCAACTGTCCCTGCCCGAGACCGGTGAGCTGGATAGCGCCACGCTGAAGGCCAT
  G  E  S  K  S  L  G  P  A  L  L  L  L  Q  K  Q  L  S  L  P  E  T  G  E  L  D  S  A  T  L  K  A  M
GCGAACCCCACGGTGCGGGGTCCCAGACCTGGGCAGATTCCAAACCTTTGAGGGCGACCTCAAGTGGCACCACCACAACATCACCTATTGGATCCAAAAC
 R  T  P  R  C  G  V  P  D  L  G  R  F  Q  T  F  E  G  D  L  K  W  H  H  H  N  I  T  Y  W  I  Q  N
TACTCGGAAGACTTGCCGCGGGCGGTGATTGACGACGCCTTTGCCCGCGCCTTCGCACTGTGGAGCGCGGTGACGCCGCTCACCTTCACTCGCGTGTACA
Y  S  E  D  L  P  R  A  V  I  D  D  A  F  A  R  A  F  A  L  W  S  A  V  T  P  L  T  F  T  R  V  Y  S
```

**Figure 1 (continued)**

```
GCCGGGACGCAGACATCGTCATCCAGTTTGGTGTCGCGGAGCACGGAGACGGGTATCCCTTCGACGGGAAGGACGGGCTCCTGGCACACGCCTTTCCTCC
  R  D  A  D  I  V  I  Q  F  G  V  A  E  H  G  D  G  Y  P  F  D  G  K  D  G  L  L  A  H  A  F  P  P
TGGCCCCGGCATTCAGGGAGACGCCCATTTCGACGATGACGAGTTGTGGTCCCTGGGCAAGGGCGTCGTGGTTCCAACTCGGTTTGGAAACGCAGATGGC
  G  P  G  I  Q  G  D  A  H  F  D  D  D  E  L  W  S  L  G  K  G  V  V  V  P  T  R  F  G  N  A  D  G
GCGGCCTGCCACTTCCCCTTCATCTTCGAGGGCCGCTCCTACTCTGCCTGCACCACCGACGGTCGCTCCGACGGCTTGCCCTGGTGCAGTACCACGGCCA
A  A  C  H  F  P  F  I  F  E  G  R  S  Y  S  A  C  T  T  D  G  R  S  D  G  L  P  W  C  S  T  T  A  N
```

A (Exon 6, +23 polymorphism)

```
ACTACGACACCGACGACCGGTTTGGCTTCTGCCCCAGCGAGAGACTCTACACCCGGGACGGCAATGCTGATGGGAAACCCTGCCAGTTTCCATTCATCTT
  Y  D  T  D  D  R  F  G  F  C  P  S  E  R  L  Y  T  R  D  G  N  A  D  G  K  P  C  Q  F  P  F  I  F
CCAAGGCCAATCCTACTCCGCCTGCACCACGGACGGTCGCTCCGACGGCTACCGCTGGTGCGCCACCACCGCCAACTACGACCGGGACAAGCTCTTCGGC
  Q  G  Q  S  Y  S  A  C  T  T  D  G  R  S  D  G  Y  R  W  C  A  T  T  A  N  Y  D  R  D  K  L  F  G
TTCTGCCCGACCCGAGCTGACTCGACGGTGATGGGGGGCAACTCGGCGGGGGAGCTGTGCGTCTTCCCCTTCACTTTCCTGGGTAAGGAGTACTCGACCT
F  C  P  T  R  A  D  S  T  V  M  G  G  N  S  A  G  E  L  C  V  F  P  F  T  F  L  G  K  E  Y  S  T  C
GTACCAGCGAGGGCCGCGGAGATGGGCGCCTCTGGTGCGCTACCACCTCGAACTTTGACAGCGACAAGAAGTGGGGCTTCTGCCCGGACCAAGGATACAG
  T  S  E  G  R  G  D  G  R  L  W  C  A  T  T  S  N  F  D  S  D  K  K  W  G  F  C  P  D  Q  G  Y  S
```

**Figure 1 (continued)**

```
TTTGTTCCTCGTGGCGGCGCATGAGTTCGGCCACGCGCTGGGCTTAGATCATTCCTCAGTGCCGGAGGCGCTCATGTACCCTATGTACCGCTTCACTGAG
 L  F  L  V  A  A  H  E  F  G  H  A  L  G  L  D  H  S  S  V  P  E  A  L  M  Y  P  M  Y  R  F  T  E
GGGCCCCCCTTGCATAAGGACGACGTGAATGGCATCCGGCACCTCTATGGTCCTCGCCCTGAACCTGAGCCACGGCCTCCAACCACCACCACACCGCAGC
G  P  P  L  H  K  D  D  V  N  G  I  R  H  L  Y  G  P  R  P  E  P  E  P  R  P  P  T  T  T  T  P  Q  P
CCACGGCTCCCCCGACGGTCTGCCCCACCGGACCCCCCACTGTCCACCCCTCAGAGCGCCCCACAGCTGGCCCCACAGGTCCCCCCTCAGCTGGCCCCAC
  T  A  P  P  T  V  C  P  T  G  P  P  T  V  H  P  S  E  R  P  T  A  G  P  T  G  P  P  S  A  G  P  T
AGGTCCCCCCACTGCTGGCCCTTCTACGGCCACTACTGTGCCTTTGAGTCCGGTGGACGATGCCTGCAACGTGAACATCTTCGACGCCATCGCGGAGATT
 G  P  P  T  A  G  P  S  T  A  T  T  V  P  L  S  P  V  D  D  A  C  N  V  N  I  F  D  A  I  A  E  I
GGGAACCAGCTGTATTTGTTCAAGGATGGGAAGTACTGGCGATTCTCTGAGGGCAGGGGGAGCCGGCCGCAGGGCCCCTTCCTTATCGCCGACAAGTGGC
G  N  Q  L  Y  L  F  K  D  G  K  Y  W  R  F  S  E  G  R  G  S  R  P  Q  G  P  F  L  I  A  D  K  W  P
CCGCGCTGCCCCGCAAGCTGGACTCGGTCTTTGAGGAGCCGCTCTCCAAGAAGCTTTTCTTCTTCTCTGGGCGCCAGGTGTGGGTGTACACAGGCGCGTC
  A  L  P  R  K  L  D  S  V  F  E  E  P  L  S  K  K  L  F  F  F  S  G  R  Q  V  W  V  Y  T  G  A  S

GGTGCTGGGCCCGAGGCGTCTGGACAAGCTGGGCCTGGGAGCCGACGTGGCCCAGGTGACCGGGGCCCTCCGGAGTGGCAGGGGGAAGATGCTGCTGTTC
 V  L  G  P  R  R  L  D  K  L  G  L  G  A  D  V  A  Q  V  T  G  A  L  R  S  G  R  G  K  M  L  L  F
AGCGGGCGGCGCCTCTGGAGGTTCGACGTGAAGGCGCAGATGGTGGATCCCCGGAGCGCCAGCGAGGTGGACCGGATGTTCCCCGGGGTGCCTTTGGACA
S  G  R  R  L  W  R  F  D  V  K  A  Q  M  V  D  P  R  S  A  S  E  V  D  R  M  F  P  G  V  P  L  D  T
```

**Figure 1 (continued)**

```
CGCACGACGTCTTCCAGTACCGAGAGAAAGCCTATTTCTGCCAGGACCGCTTCTACTGGCGCGTGAGTTCCCGGAGTGAGTTGAACCAGGTGGACCAAGT
  H  D  V  F  Q  Y  R  E  K  A  Y  F  C  Q  D  R  F  Y  W  R  V  S  S  R  S  E  L  N  Q  V  D  Q  V
GGGCTACGTGACCTATGACATCCTGCAGTGCCCTGAGGACTAGGGCTCCCGTCCTGCTTTGCAGTGCCATGTAAATCCCCACTGGGACCAACCCTGGGGA
 G  Y  V  T  Y  D  I  L  Q  C  P  E  D  *

AGGAGCCAGTTTGCCGGATACAAACTGGTATTCTGTTCTGGAGGAAAGGGAGGAGTGGAGGTGGGCTGGGCCCTCTCTTCTCACCTTTGTTTTTTGTTGG
AGTGTTTCTAATAAACTTGGATTCTCTAACCTTT
```

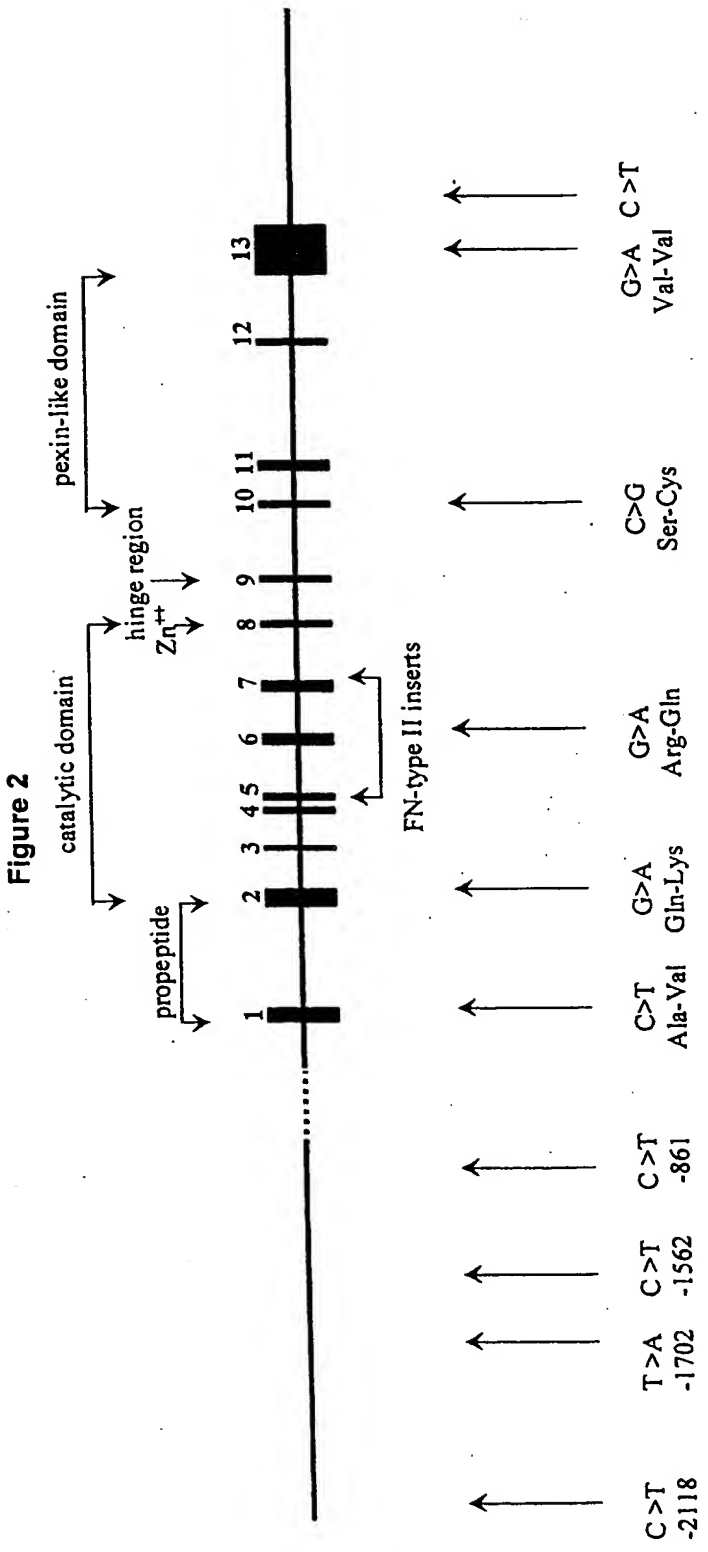
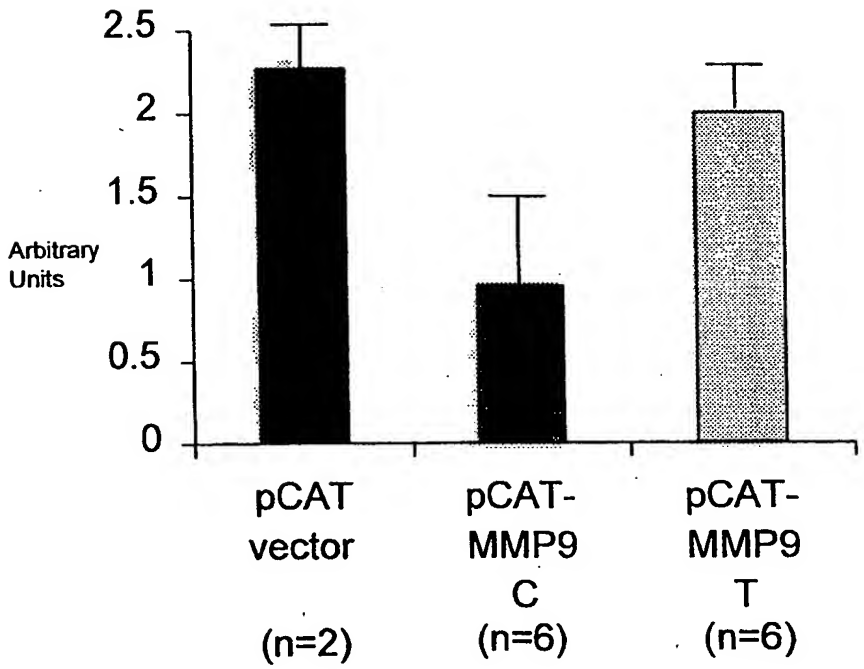

pexin-like domain
hinge region
$Zn^{++}$
catalytic domain
propeptide
FN-type II inserts
1
2
3
4
5
6
7
8
9
10
11
12
13
C>T -2118
T>A -1702
C>T -1562
C>T -861
C>T Ala-Val
G>A Gln-Lys
G>A Arg-Gln
C>G Ser-Cys
G>A C>T Val-Val

Figure 2

**Figure 3**


2.5
2
1.5
1
0.5
0
Arbitrary Units
pCAT vector (n=2)
pCAT-MMP9 C (n=6)
pCAT-MMP9 T (n=6)